<u>AFFIDAVIT OF DANIEL KRIEGMAN, Ph.D.</u>

I, Daniel Kriegman do solemnly swear under the pains and penalties of perjury the

following is true to the best of my knowledge.

Professional Background

1.   I have been a licensed psychologist in Massachusetts since 1981. I received my B.A.,

with a major in Psychology, from State University of Buffalo. I received my M.A. and Ph.D. in

Clinical Psychology from Boston University. I am a member of various professional associations.

I have also been a designated Qualified Examiner under M.G.L. c. 123A, responsible for the

evaluation of sexual dangerousness for commitment hearings in Massachusetts. My Curriculum

Vitae is attached as Exhibit 1.

2.   I have significant experience in the area of sex offender evaluation and treatment.

Between 1977 and 1986, I was a staff Psychologist, Principle Psychologist, Chief Psychologist,

and eventually a Super Chief Psychologist, at the Massachusetts Treatment Center for the

Sexually Dangerous. My responsibilities included providing both individual and group

psychotherapy, psychodiagnostic testing and preparing background summaries and clinical

formulations used for the determination of sexual dangerousness of men who were civilly

committed under the sexual dangerousness statute in Massachusetts. In 1982, I was named Unit

Director for one of the two units at the Treatment Center and thus supervised all the clinical staff,

treating approximately 130 sexually dangerous persons on the unit. Eventually, I became the

Director of Supervision and Training responsible for the supervision and training of all the

clinical staff at the Treatment Center. I was simultaneously the Director of Intake and Treatment

Planning, responsible for evaluating inmates for possible commitment and for formulating

individualized treatment plans if the offender was committed.

3.   After I left the Treatment Center in 1986, I was designated a Qualified Examiner. In the

late 1980's, I founded Human Services Cooperative, Inc. (HSC), which contracted with various

agencies to provide psychological treatment and evaluation. Around 1989, the Commonwealth of

Massachusetts placed all sexual dangerousness determinations (screenings in prison, initial

evaluations during a 60-day commitment to determine if a man was sexually dangerous, and

periodic reviews to determine if men who had been determined to be sexually dangerous were

still so) in a contract and sought a single vendor to provide all sexual dangerousness opinions.

HSC won the bid and for several years, under my supervision as President of the Corporation and

Clinical Director, provided all of the sexual dangerousness determinations for the

Commonwealth. HSC also provided all independent (non-state employee) psychological and

psychiatric services at the Treatment Center under a contract with the Department of Mental

Health (which ran the institution at the time). These latter services included the two independent

psychiatrists or psychologists mandated by the statute to sit on the five member Community

Access Board, which performed annual reviews of the treatment and sexual dangerousness status

of all the sexually dangerous men in the Commonwealth.

4.   Since leaving the Treatment Center in 1986, I have continued to provide psychotherapy to

sex offenders. I have also continued to evaluate persons for sexual dangerousness in court

proceedings and, since 2002, for the Sex Offender Registry Board. In the three decades (1977 to

2007) during which I have worked with and evaluated sex offenders, I have personally completed

well over 400 sexual dangerousness evaluations of sex offenders, have participated on a team

involved in such evaluations on several hundred additional occasions, and have overseen well over 1,000 additional evaluations.

5.   I have authored or co-authored or co-edited over 30 published articles and books in the field of psychology including articles specifically related to human sexuality and sex offender treatment. I am on the Editorial Board for *The Journal of Sexual Offender Civil Commitment*. I have been a teaching fellow, led various seminars, delivered papers at international conferences and taught at select institutions.

6.   In total, I have personally treated over 100 sex offenders in the course of my career. I have also evaluated or participated in the evaluation of over 1,000 sex offenders for dangerousness in the course of my career. I have testified as an expert in court on over 350 occasions, both for the Commonwealth and for the defense.

7.   I have an extremely thorough understanding of the scientific literature regarding sex offender treatment and evaluation. As an undergraduate, I was a teaching assistant leading a section for the required psychology course "Research Methods," (the only time an undergraduate ever performed such a function at SUNY Buffalo, where I graduated Magna Cum Laude with a B.A. in Psychology). I have been a teaching assistant at the graduate school level and have tutored Ph.D. candidates in statistics and research methods. I have served as a peer review editor for several scholarly journals in psychology.

8.   My experience and personal knowledge are based on the Massachusetts commitment scheme for sexually dangerous persons, one of the oldest commitment laws in the nation. When relevant, I will refer to pertinent experience from Massachusetts as a way to explain the problems associated with predicting dangerous behavior and determining sexual dangerousness.

Early History About the Evaluation of Sex Offenders

9.   Psychologists have long understood the limitations of predicting human behavior. Though we have been called on to opine whether someone is dangerous, and in recent years, sexually dangerous, historically, the profession had come to consider it impossible to make such predictions with a substantial degree of accuracy. Thus, the main professional associations (the American Psychological Association and the American Psychiatric Association) have determined that offering such clinical expertise in the courtroom is often unethical unless the clinician makes the very limited validity of such opinions clear. Since this rarely occurs, the main professional associations for psychiatrists and psychologists (the APA's) have questioned the ethics of offering such "expertise" in the courtroom.

10.   Up until the 1980's, the predominant tool in use for making predictions of future sex offenses (or other forms of violent, dangerous behavior) was clinical judgment. Clinical judgment is based on a clinician's subjective sense of an offender when the specifics of the offender's life and the experience of interacting with the offender is "mulled over" in the mind of a clinician who has had experience working with other sex offenders. However, the reliability of clinical judgment was always in doubt and when empirically tested by seeing if clinicians could accurately predict which offenders will reoffend, the results were, on average, approximately equal to tossing a coin.

11.   Professor Meehl was one of the early explorers attempting to determine the accuracy of clinical prediction. His work provided the foundation for later understanding of this issue.[1]

---

[1]A bibliography listing the articles referenced in this Affidavit and the attached Exhibits is included as Exhibit 12.

Around the same time, an early classic study by Professor Goldberg demonstrated that novices were often able to formulate clinical opinions with equal accuracy to those formulated by seasoned professionals. This finding has been replicated many times.

12.   By 1981, Professor Monahan published his seminal work establishing that clinical predictions of dangerousness were wrong 2 out of 3 times; clinicians are often poorer predictors than lay people. I have provided a more detailed historical background in Exhibit 2 (History of Clinical Prediction Studies).

Use of Actuarials

13.   The early research about clinical judgment spurred new fields of research into the effectiveness, and creation of, alternative methods of prediction. Actuarials have long been used in other contexts, such as insurance. In the 1980's, psychologists began researching and developing actuarials in the field of violence prediction.

14.   Before getting into the history of actuarials, it is useful to understand how they work and what their limitations are.

15.   Actuarial tests—of which life insurance actuarial methods are excellent examples—take a number of predictive factors that are each known to be independently and truly related to, for example, longevity. Age, weight, smoking behavior at time of application for insurance, among other factors, are each assessed and scored according to the magnitude of each. The different factors are *weighted* to emphasize more powerful predictors (e.g., age) over weaker predictors. And then the factors are combined according to an actuarial formula to produce a prediction. For example, all other things being equal, a sixty-year-old man has a shorter life expectancy than a

thirty-year-old smoker; age would influence the risk estimate more than smoking. However, a

72-year-old smoker would have a shorter life expectancy than a 75-year-old nonsmoker; in this

case, smoking would exert more influence than the age difference. The *actuarial formula*

combines and weights the various valid predictors to produce predictions that maximize accuracy

in a large, real-life sample. The formula is "tweaked" and reworked until predictive accuracy is

maximized. It is then cross-validated on new samples and further modifications in the actuarial

formula are made to increase the predictive accuracy in a wide range of relevant samples.

16.   There are several important statistical components to actuarials which allow a

psychologist to apply them in the appropriate contexts. It is important to have a rudimentary

understanding of the various components of an actuarial in order to appreciate its usefulness and

limitations.

17.   The relevance of all actuarials—indeed, all opinions regarding sexual

dangerousness—begins with a base rate. No facts about an individual (person or event) can have

any meaning unless they are compared to the class of similar facts about similar individuals and

events. That is, it is very rare that the predictive meaning of an event can be interpreted without

knowledge of base rates. Base rates are, very simply, a reference to the class of other similar

events.

18.   Base rates are important for two reasons. First, the lower the base rate, the harder it is to

be accurate when concluding that the target behavior is likely (in this case, future sex offending)

and the more likely that false positives (false conclusions of dangerousness) will be produced

than accurate "hits." If the percentage of convicted sex offenders released after attaining the age

of fifty who recidivate is 5%, it will be extremely hard to predict accurately which persons over

the age of fifty will recidivate (correct predictions of dangerousness) without including a larger number of non-recidivators (false positives); on the other hand, if the percentage of persons who have multiple convictions for sex offenses and are under the age of 25 who recidivate is over 50%, it will be far easier to be right more often than not when making the prediction that a specific offender from this group will reoffend.

19.   Additionally, base rates are important because they help determine when and how to apply a specific risk assessment to a given case. For example, if a base rate were obtained by studying offenders with a prior conviction, over a certain age, and with a male victim, such a base rate would be expected to produce accurate results when applied to offenders with the same characteristics. To the degree that an offender does not match the group on which a base rate was determined, that base rate would be misleading if applied to the offender in question.

20.   Actuarials are able to calculate base rates by combining and calculating a number of significant variables in a manner that increases the accuracy that could be obtained by using any one factor. For a more complete discussion and explanation about base rates, see Exhibit 3 (Base Rates).

21.   The highest recidivism estimate ever produced based on empirical observations of what actually occurs when sex offenders are released and followed for 25 years (approximately 50% for a certain category of offenders) comes from Prentky, et al. (1997). The Prentky study was based on men who had been adjudicated sexually dangerous in Massachusetts and subsequently released from prison. See Exhibit 4 (Prentky Study). This study, however, is likely to be an overestimate of recidivism for a number of reasons.

22.   It is more commonly believed that the base rate for sex offender recidivism as a whole

lies between 15 and 35%. This base rate fluctuates depending on various factors. For example,

studies consistently confirm that older offenders have a significantly lower base rate; on the other

hand, offenders under 25 have a higher base rate. Extra-familial child molesters have a higher

rate of recidivism than men whose victims were adults. Both of these groups have a higher rate

of recidivism than incest offenders, who have the lowest rate of recidivism in this simple

classification schema. In all, sex offenses have been shown to have the *lowest* rate of recidivism

of all common offense types (only murderers recidivate at a substantially lower rate).

23.   There are other aspects of actuarials that are important in understanding their limitations.

Actuarials are able to tell us how well a given set of traits or characteristics correlate with the

studied outcome, in our case, recidivism. The correlation can be measured by a correlation

coefficient, (denoted by r). Consider that the absolute value of r can range from 0.0 to 1.0, with

1.0 being a "perfect" relationship and 0.0 being no relationship at all, i.e., 0.0 is the equivalent of

what one would get by tossing a coin and saying "Heads, he'll reoffend; Tails, he won't." That

is, if the correlation between two variables yields an $r = 1$, then knowing how an individual

scores on one variable tells you perfectly how they will score on the other.

24. Correlation coefficients between 0.0 and 1.0 have different predictive values. A correlation

coefficient less than .l0, i.e., an r close to zero, has little predictive value and virtually no utility

in practical applications. A correlation coefficient less than .20 has such a small amount of

predictive value that, in a real life application in a small number of cases, it would be hard to see

a difference between tossing a coin and using such a predictor.   Between .20 and .40 there is an

increase in accuracy (over chance or coin tossing) that would be perceptible. From .40 to .70

there are clear advantages to using the predictor. Above .70 we have a fairly strong predictor that is far better than random guessing.

25.   By combining predictive factors with correlations between .10 and .25, it is possible to create actuarial tests that produce correlations between .20 and .40 with sex offense recidivism. Experts agree that this type of actuarial prediction has been conclusively shown to be the essence of the most valid predictive methods currently available; nevertheless, this actuarial prediction is not very strong (valid). Given the aforementioned problem of making predictions about phenomena with low base rates, it becomes vital to understand the real but limited validity of our best predictive method in order for a court to be able to decide if this is sufficiently valid to be used in a civil commitment proceeding. For a more detailed discussion of correlation coefficients, see Exhibit 5 (Correlation Coefficients, Validity, and Variance).

26.   Given large enough samples, it is possible to show that weak correlations with no practical utility are real, and that therefore the small correlation is unlikely to be just a random deviation from chance. This means that if we examine a different large sample, we can expect to find the correlation again and again. However, when comparing small groups of individuals or single individuals with or without the weakly correlated factor, there will be no discernable difference between the presence or the absence of the factor. That is, with small groups (a small sample), we won't be able to perceive the increase in accuracy obtained by using a weakly correlated factor (e.g., $r = .10$) over the random variations we will produce when tossing a coin ($r = 0.0$) a small number of times. Thus, simply knowing that a given trait has a significant (or real) correlation does not, in and of itself, tell us how much more accurate a prediction will be when the presence (or absence) of that trait is utilized in making a prediction.

27. However, given the correlation, there is a way of determining how *valid* a prediction is. Validity means how much we can count on knowledge of one variable (clinical or statistical assessment of risk) to tell us what in fact will happen when we look at another variable, in this case actual recidivism. This is known as the percent of variance (variation in outcome) in an outcome predicted by a variable predictor . For a more detailed discussion of validity and variance, see Exhibit 5.

28. The leading researcher in the field of sex offender recidivism is Karl Hanson. In 1998, he published a meta-analysis of all known sex offender recidivism studies. That meta-analysis (and a follow-up published in 2004) became the foundation for modern risk assessment in the field of sex offense recidivism. In those analyses, factors that had been empirically demonstrated to be correlated with recidivism were identified. Numerous factors that had been *thought* to be correlated but which are not predictive were also identified. Based on this information, Dr. Hanson and a colleague developed an actuarial tool called the Static-99, which was based on formulas taking into account various characteristics (factors), each of which had been shown to demonstrate some predictive value.

29. Since the development of the Static-99, Dr. Hanson has also led the way in studying the effects of age and recidivism, demonstrating that age seriously decreases the rate at which sex offenders reoffend.

30. The development of actuarials has changed the way we think about risk prediction. In its increased accuracy over clinical opinion, the actuarials highlight how unreliable clinical judgment is; yet, the actuarials produce only limited validity of their own. Actuarials help demonstrate that in the real world of sex offense recidivism prediction, the best predictive

method has resulted in only a small degree of validity. Furthermore, this greater validity over

clinical opinion highlights how very little correlation there is between an expert's clinical opinion

that someone is sexually dangerous and whether that person will actually recidivate. For a more

detailed discussion of the specific application of the broad statistical concepts to these hearings,

see Exhibit 6 (Specific Application of Statistical Concepts to Findings of Sexual Dangerous).

31.   Actuarials also reveal the extent of their own limitations. Although it is true that the

actuarial method produces results that are *much* more accurate than clinical judgment (indeed,

the best actuarials may be more than ten times more accurate than clinical judgment), the

predictive power is far from perfect. A perfect predictor would account for 100% of the

"variance" in an outcome (in this case, the "variation" is reoffend or does not reoffend). A

useless predictor, e.g., tossing a coin, will account for 0% of the variance. Clinical judgment

accounts for less than ½ of 1% of the variance, making it almost useless. If the actuarials can

predict 10 to 20 times more accurately than clinical judgment (which they can), this does *not*

make them highly accurate. In fact, they have been shown to have a correlation of between .3

and .35, thus predicting just 9 - 12% of the variance, leaving about 90% unaccounted for.   For a

more detailed example, see Exhibit 7 (Application of Actuarial Correlations).

32.   Today, variations of clinical judgment and the use of pure actuarials have evolved. Some

professionals use a method called guided clinical, which calls for the clinician to evaluate a list

of factors with valid (real) small correlations with recidivism. The clinician then uses clinical

judgment to combine the information from the list of factors into a single risk estimate. This

method has demonstrated validity between the use of pure, unstructured (unguided) clinical

judgment and the use of actuarial prediction. Others use a method called the adjusted actuarial

method, which uses actuarials to establish a base rate for a typical group of offenders similar to

the offender in question. The clinician is then free to adjust this risk estimate by using clinical

judgment to raise or lower the risk estimate when unvalidated factors (factors believed to be

correlated with recidivism that have not been researched sufficiently to determine if they are so

correlated) or other validated factors not taken into account by the actuarial formula are present.

However, none of these have proven to be substantially more accurate than pure clinical

judgment and limited research indicates that they are not as accurate as the pure actuarial method.

Since the actuarial method has limited accuracy, these weaker methods cannot produce the level

of accuracy needed for confident conclusions and the use of these alternate methods has not

gained acceptance in the psychological field. For a more detailed explanation of these variations,

see Exhibit 8 (Clinical Variations). It still remains the case that simply using an actuarial upon

which to base an opinion of dangerousness is the most accurate of the various methods of

prediction, none of which is highly accurate.

33.   Since the accuracy of our most accurate methods of prediction is poor, we can

rarely have a sufficient degree of professional certainty that future offenses are likely to

occur when: 1) the base rate is unknown but likely to be fairly low, between 15 and 35%,

2) we know that the base rate does not predict very reliably in many cases, 3) the factors

we can use to modify the base rate are weak predictors with fairly low validity, and 4)

many years have passed since the last sexual offense. See Exhibit 9 (The Psychological

Limitations of Opining Sexual Dangerousness) and Exhibit 10 (Why an Actuarial

Prediction Can Never Be a Sufficient Basis for a Finding of Sexual Dangerousness).

Jimmy Ryce Civil Commitment Program

34.   Under the Program, a person declared sexually dangerous may be released if the "Director of the facility in which a person is placed pursuant to subsection (d) determines that the person's condition is such that he is no longer sexually dangerous to others, or will not be sexually dangerous to others if released under a prescribed regimen of medical, psychiatric, or psychological care or treatment." This presents several problems already addressed above. For one, if psychologists themselves cannot accurately predict whether someone will recidivate, and indeed over-predict recidivism more often than not, there is no reason to believe a Director has an accurate predictive ability.

35.   My experience with the Massachusetts commitment scheme is that the Department of Corrections has never filed a petition for release and opposes an inmate's petition for release approximately 99% of the time, even after offenders have been civilly committed and treated for over a decade. Therefore, the constraints inherent in a psychological prediction are more apparent when the Government is responsible for declaring someone is no longer dangerous (as opposed to a state expert declaring someone is). For a more detailed discussion about an inmate's release in Massachusetts, see Exhibit 11 (Release from Commitment in Massachusetts).

Conclusion

36.   In conclusion, in three decades of working with sex offenders, evaluating them for sexual dangerousness, and studying the research—the quantity of which has exploded in the last 15 years—it has become clear that we simply have no ability to accurately predict sex offense recidivism for individuals.

Date: May 16, 2007   /s/ Dr. Daniel Kriegman

                                                        Dr. Daniel Kriegman